



Training and Educational Evaluation

A Journal for Marine Corps Trainers and Educators



June 2004

Quarterly Electronic Journal

Volume 1, Issue 1

Table of Contents

Editor's Note.....	1
IMS East.....	1
IMS West.....	2
Important Dates.....	3
Training Conferences.....	3
Teaching Strategies and Tips.....	4
Picture Memory Strategy For Students and Instructors by Jeffrey M. Harrington, Ed.D.	
Evaluation Tips.....	6
Instructional System Development – Evaluation Phase by Donald Clark, copyright 1995, revised 1996 and 1997	
Student Evaluation of Instruction – A Case for Validity of Instructional Rating Forms.....	9
by Christina R. Silence	

Schools and Training Centers where over 80,000 Marines are trained—that's over one-third of the Marine Corps annually!

In training and educating our Marines and service members we assume a number of roles as professionals. By the use of the term roles, I am speaking of the broad areas of responsibility within workplace learning and performance that require a select group of competencies and areas of expertise to perform effectively. Roles are not the same as job titles; they are much more fluid depending on the work or project. For the formal school instructor, curriculum developer, and administrator, playing different roles is analogous to maintaining a collection of hats: When the situation calls for it, the practitioner takes off one hat and dons another.

It is the intent of this journal to provide you with information to successfully serve in all of your roles.

Semper Fidelis,

Capt J.M. Harrington

Instructional Management School East



Director, IMS East

On behalf of the staff of Instructional Management School (East) I would like to welcome you to the first edition of our journal. In this journal and in conjunction with our peers at IMS (West) we hope to bring to you a quarterly snapshot of where we are and where we are going with formal school education through the eyes of the IMS'.

IMS is evolving. Not only have we developed new courses based on the new (and soon to be signed) Systems Approach to Training Manual and the recently signed 1553.2A, but the face of IMS (East) has changed dramatically with an almost 100% turnover in personnel in less than a year's time. Fortunately, our replacement personnel all come from a solid background in formal school education so the transformation was seamless to our students. Additionally, we just recently hosted the IMS Director's Conference here at IMS East which was attended by the staffs of both IMS' as well as by our Task Analyst from TECOM, Capt Harrington. We discussed various issues ranging from throughput to SOP and ultimately made a lot of progress in streamlining the way the IMS' support the needs of their customers.

Marines, Sailors, and Civilian Marines,

Welcome to the inaugural edition of ***Training and Educational Evaluation***. The purpose of this journal is to serve two functions: first, to serve as a conduit of Marine Corps Training Information through the Instructional Management Schools East and West; and, second, to serve as a professional journal of Marine Corps and military educators and trainers. To this end, you will see three main sections of this journal: a section dedicated to Instructional Management School East, a section dedicated to Instructional Management School West, and a section dedicated to training and educational information for administrators, curriculum developers, and formal school instructors.

To this end, you will see training tips and scholarly articles that will provide the trainer and educator with information applicable to their trade and craft. In this regard, we will be soliciting letters to the editor and will also provide a forum for publishing articles. To maintain a level of professional credibility all articles submitted for publishing will be vetting through a formal process. Letters and articles may be submitted to the Director of Instructional Management School East and West or to the following email address: harringtonjm@tecom.usmc.mil

If you are wondering whether or not you are a professional trainer or educator consider this fact: personnel assigned to formal schools within the Training and Education Command are responsible for monitoring, managing, and developing Individual Training Standards (ITSs) for students, Programs of Instruction, and providing evaluation at all Formal

We are your resource! If you have ideas, questions, or just want to run an idea by somebody, give us a call. Not only are we here to train your formal school instructors, curriculum developers and school administrators, we truly enjoy assisting the formal schools in making their establishments more efficient in meeting the needs of their students. You can visit our website at [MCCSSS Website](#), click on MCCSSS Schools, then Instructional Management. There you will find our current course schedule, course descriptions as well as emails to various points of contact at IMS. Please let us know if we can be of assistance to you in any way.

Semper Fidelis,

Mike Milburn
Capt USMC



SNCOIC, IMS East

With the warmer weather descending upon us, a couple of issues come to mind. First and foremost is the seasonal uniform change. As of Monday, 29 March 2004, the uniform of the day will be the utility uniform with sleeves up and Service "C" becomes the graduation uniform. The other issue is hydration. Make sure that you are drinking plenty of water everyday and start thinking about adjusting the time of day in which you conduct PT.

As formal schools become more familiar with MCO 1553.2A, there is a need to address the prerequisites. All personnel attending the appropriate Instructional Management School (IMS) course for their billet are required to complete the System Approach to Training (SAT) and the Operational Risk Management (ORM) Distance Learning (DL) courses prior to attending IMS. These courses are located on MARINET <http://www.marinenet.com>. Each prospective student needs to print their graduation certificates and bring the certificates with them when they report to IMS. Personnel attending the Formal School Instructor Course (FSIC) are also required to bring one completed Master Lesson File (MLF) that contains a lesson that is 50 minutes in length with lecture, demonstration and Practical Application as it's methods. It is very important that the lesson be as close as possible to 50 minutes because longer lessons take a lot of time to adjust and it's an added burden that the student does not need in this fast paced course. Personnel who signed up for the Curriculum Developer Course (CDC) are required to complete the Marine Corps Automated Instructional Management System (MCAIMS) tutorial. The tutorial is located on the Ground Training Branch Homepage <http://www.tecom.usmc.mil/gtb/MCAIMS/>

Last but not least, IMS is here to support the Formal Schools/Detachments. Besides our resident course we send out Mobile Training Teams (MTT) and I am currently looking at FY05 and FY06 schedules. Also we conduct Curriculum Assist Visits (CAV) to assist schools at all levels. If you would like an

MTT or CAV just give us a call at Comm (910) 450-0898 DSN 750-0898 or e-mail me at david.monaghan@usmc.mil

D. T. Monaghan
MSgt USMC



Academics Officer, IMS East

It's amazing how much information that we have easy access to on the internet today. It only takes a few keystrokes here and there, and "bingo", your own personal library is right in front of you. Whether you're an instructor, curriculum developer, and/or administrator, there are some great websites available to assist you in your effectiveness. We all have an opportunity to dig deeper and learn more about the many facets related to training. Below are a few websites that may be of interest.

Instructional Management School West



Director, IMS West

I am Captain Tony Maldonado and have officially been the Director of the IMS since February 2004. I have been non stop ever since. My goal during my tenure is to ensure that all selected personnel assigned to supervisor and instructional billets in Marine Corps schools, the Fleet Marine Force, Marine Corps Reserve, and other services are trained in the foundation of design, execution, and management of systematic instruction. We at IMS West are committed to provide all of our students with a quality educational experience. The training schedule for 2004 consists of a total of 624 school seats spread across all three courses: FSIC, CDC, and AC, which are being taught in 34 iterations. As of February 2004, the FSIC graduated 168 students (excluding MCIC FSIC graduates), the CDC course graduated 33, and the AC graduated 28 students. We are now quickly approaching the third quarter and our enrollment is steady. We are facing an instructor shortage at the school, but I am optimistic that we will be sufficiently manned in time for the fiscal year 2005 school year. Again, I am proud to be serving in such a fine institution and I am eager to provide quality education to all of our forces. My number is DSN: 365-4271 or commercial (760) 725-4271.



SNCOIC, IMS West

I am MGySgt Garcia and the SNCOIC of IMS West. I feel it is my responsibility to support this school in all matters relating to academic, student and instructor welfare, while fostering an atmosphere conducive to learning for both the

student and instructors alike. Current concerns are instructor replacements, which have proven to be quite challenging. We started an all out recruiting blitz, which has yielded one inbound GySgt and a list of other potential candidates. Advertisement of this school has been and will continue being one of my main concerns to ensure that our academic community as well as any other interested parties has a complete knowledge of our mission and capabilities. Other concerns include maintenance of current instructional equipment in particular existing warranties and responsibilities of repairs. Finally, the NMCI introduction is a concern due to our heavy reliance on computer equipment for use by the students. Other than that training and management of our instructor staff is a continuous process to include future planning for equipment and building upgrades.



Instructional Systems Specialist, IMS West

I am Mr. Michael Rank, the Instructional System Specialist for IMS West. I teach in the Administrator and Curriculum Developer courses. Currently, I have been working on incorporating a new set of performance examinations from the Curriculum Developer Course into MCAIMS.



Instructional Systems Specialist & Administrator Course Chief, IMS West

Currently, the Administrator Course is three days in length and focus' on many of the problem areas new administrator's experience within the formal school. Administrators learn how the Systems Approach to Training (SAT) and MCO 1553.2A affect their school resources. They learn how to gather and analyze evaluation data to determine the effectiveness of their instruction and how to identify problem areas. Staff and faculty development are discussed with an eye toward maximizing limited resources to produce a plan that trains personnel at all levels of the formal school. Finally, students participate as members of a Course Content Review Board (CCRB) to recommend modifications to courses based on collected data. Everyone that attends the Administrator Course leaves with new ideas of how they can improve their schools. Their interaction with other administrators allows them to acquire new perspectives on problem resolution that will benefit their schools. The return on the investment (three days) will equip the administrator with skills not available from any other resource and is time well spent. The following are the remaining AC coursed for 2004:

M10rf91	2004030	30 Aug – 3 Sep 04
M10rf9m	2004020	26 Jul – 30 July 04



FSIC Chief, IMS West

I am GySgt. M. Juarez and my mission is to build a foundation from which an instructor can perform the competencies expected of him/her by the Marine Corps. So how is this accomplished? I believe that this is done by training the instructor to prepare and employ a period of instruction. The instructor also learns the skill sets necessary to Administer student test. Below is a list of classes left for FY 2004.

IMS Camp Pendleton:

Class ID	Dates
2004030	5/17/04 – 5/28/04
2004040	8/02/04 – 8/13/04

Mobile Training Team

2004050	4/05/04 – 4/16/04
2004055	5/03/04 – 5/16/04
2004060	6/21/04 – 7/02/04
2004070	8/02/04 – 8/13/04
2004080	9/13/ 04 – 9/24/04

Marine Corp Instructor Program

2004055	4/21/04 – 4/30/04
2004065	7/14/04 – 7/23/04



CDC Chief, IMS West

The Curriculum Developer Course is a two-week course focusing on the design and development of performance-based curriculum. Currently, the CDC is taught by SSgt. Halstead (Course Chief), Mr. Rank (Instructional Systems Specialist), and Mr. Hays (Instructional Specialist). Throughout the course students are required to develop a Program of Instruction (POI) from inception to completion. They begin with an Individual Training Standards (ITS) Order and finish the course with everything required to implement instruction (Lesson Plans, Media, Testing, etc.) Currently there are only a few more classes being taught during the remainder of this fiscal year. The course convening dates for the remaining resident courses are 20040621 and 20040913, and for the Mobile Training Teams to Twenty-nine Palms are 20040607 and 20040816. For more information or to inquire about a school seat email IMS@pendleton.usmc.mil.



Upcoming Training Conferences

**42nd Annual International Performance
Improvement Conference and Exposition**

Tampa Marriott Waterside Hotel & Tampa
18-23 April 2004
<http://www.ispi.org/home>

2004 Performance-Based Instructional Systems Design (ISD) Conference

Focusing on Results
Chicago Illinois
Thursday, September 30 to Saturday, October 2
<http://www.ispi.org/home>

Interservice/Industry Training, Simulation & Education Conference

Orange County Convention Center, Orlando, Florida
6-9 December 2004
<http://www.iitsec.org/>

Bob Pike's Train-The-Trainer Boot Camp

21-22 June 2004	Minneapolis, MN
27-28 July 2004	Minneapolis, MN
17-18 November 2004	Phoenix, AZ
6-7 December 2004	New York, NY

<http://www.bobpikegroup.com/seminars/>



Teaching Strategies and Tips

PICTURE MEMORY STRATEGY FOR STUDENTS AND INSTRUCTORS

By
Dr. Jeffrey M. Harrington, Ed.D.

If you stop for a moment and try to imagine life without memory, our cognitive system (or systems) for storing and retrieving information, you'll see at once that it is truly a crucial aspect of our cognition. If we did not possess memory, we would be unable to remember the past, retain new information, solve problems, or plan for the future. When are we most aware of memory? Typically, when it fails—when we are unable to remember information that we need at a particular moment. Often, memory seems to let us down just when we need it most—for instance, during an exam or some type of evaluation. Why does this occur? Why is information entered into long-term memory sometimes lost, at least in part, with the passage of time? Many explanations have been offered, so here we'll focus on the ones that have received the most attention.

The earliest view of forgetting was that information entered into long-term memory fades or decays with the passage of time. While this seems to fit with our subjective

experience, many studies indicate that the amount of forgetting is not simply a function of how much time has elapsed; rather, what happens during that period of time is crucial (e.g., Jenkins & Dallenbach, 1924). So, early on, psychologists rejected the notion that forgetting stems from passive decay of memories over time and turned instead to the following views: retroactive interference and proactive interference. Retroactive interference is simply information currently being learned interferes with information already present in memory. For instance, if learning how to operate a new computer program causes you to forget how to operate one you learned previously, this is an example of retroactive interference. In proactive interference, in contrast, previously learned information present in long-term memory interferes with information you are learning at present. For instance, suppose you learned how to operate one VCR; now you buy a new one, which requires different steps for recording a television program. If you now make mistakes by trying to operate the new VCR in the same way as you did the old one, this constitutes proactive interference. You may have probably experienced both proactive and retroactive interference with updates to new curriculum and equipment.

Let's turn our discussion of the different kinds of information stored in memory. One important type involves factual information. Memory for such information is sometimes termed explicit or declarative memory, because we can bring it into consciousness and report it verbally. It consists of two major types: episodic memory and semantic memory. Episodic memory holds information we acquired at a specific time and place it is the kind of memory that allows you to go back in time and to remember specific thoughts or experiences you had in the past. Semantic memory, in contrast, holds information of a more general nature—information we do not remember acquiring at a specific time or place. Such memory includes the meaning of words, the properties of objects, typical events in everyday life, and the countless facts we all learn during our school years.

As a student going through a program of instruction or an instructor teaching a program of instruction you have lots of first hand experience with the functioning of episodic memory. Often, you must commit to memory lists of definitions, terms, or formulas. What can you do to improve such memory? Research on semantic memory suggests that many factors influence it, but that among these the most important are the amount and spacing of practice. The first finding seems fairly obvious; the more often we practice information, the more of it we can retain. However, the major gains occur at first, and then further improvements in memory slow down. For this reason, spacing (or distribution) of practice is important too. Spreading out your efforts to memorize information over time is helpful. For instance, two sessions of thirty minutes are often better, in terms of retaining information, than one session of sixty minutes. This suggests that memories somehow consolidate or grow stronger with the passage of time. Another factor that has a powerful effect on retention is the kind of processing we perform. When we study a list or words, we can simply read them or listen to them; or alternatively, we can think about them in various ways. As you probably know from your own studying, it is possible to read the same pages in a text over and over again without remembering much of the information they contain. However, if you actively think about the material and try to understand it you stand a better chance

of remembering it during the testing or evaluation phase of your educational or training experience.

Thus far, we have discussed traditional views of memory retention; that is to say, the traditional strategies for processing information into long-term memory. These strategies revolve around visual text learning and storing visual text information. **However, what if we could store the same piece of information in two places?** Logic would suggest that we would double our probability of retrieving this information from our memory. By using a picture memory strategy you can have a mental photograph of the information as well as episodic memory of the visual text.

Once an individual experiences difficulty learning material such as math facts, that require rote memorization, they cannot and should not be made to repeat that which causes them difficulty. Doing the same thing again and again and expecting different results will only serve to frustrate an individual and have them view learning as unpleasant. The most obvious solution is to change our instructional methods and help them learn the information in a way that more closely matches how they learn and how they will be tested. During the learning process, there are actually many functions going on at once; however, the three that we are most concerned with are input (how the individual take the information into the brain); storage (how the individual stores the information in the brain); and output (how the individual retrieves and utilizes the information).

When people recall visual images or pictures, their eyes move up to the left or to the right. You can determine the visual memory location by asking questions that require the person to access a picture from their memory. You need to determine whether the individual is looking up to their left or up to their right when remembering a picture. You can do this by asking the individual several questions while observing their eye movements. It is important that you continue asking questions until the individual looks up to the left or up to the right. (Looking downward or to the side places the individual in a "physical, feeling" or "auditory, hearing" learning style and will not work for learning and recalling information pictures.)

Use the following examples so that the individual must get a picture in their mind in order to answer the question. Make sure that you do not tell the individual why you are asking the questions ahead of time. You want their natural response now and later on you can tell them why you asked these particular types of questions. Make sure that you remember too, that as you look at your individual their right or left is opposite of yours. You will only record the upper right or upper left direction that is theirs, not yours.

Sample questions:

Get a picture in your mind of your house. How many windows can you count?

Where were you when you learned of the terrorist attacks on 9/11? Depending upon the response ask questions related to the place where they were located so that they will retrieve the picture in their mind.

Get a picture of your best friend or spouse in your mind. Tell me exactly how he or she looks. What color is his/her hair, eyes, etc. What does one of his or her favorite shirts look like?

These are some basic questions that you can use, but also know that if you do not get the required response from the person you are questioning continue with the same subject area but in more detail. Remember, visual memory eye movements are not the same for every person. You must determine whether it is up to the left or up to the right. (It's easier to do this when you do not tell the person ahead of time what you are looking for.) Now, when you teach anything that requires visual memory (i.e. writing on the blackboard, briefing maps, etc.), you can actually use that eye movement information to physically place the word or information to be learned either up to the left or up to the right, so that retrieval is consistently successful. You will actually be showing the student how to access visual memory by having them use their eye movements to see a picture in their mind. Using different eye movement positions will actually change the learning style.

This visual memory will also work when studying tech manuals or visual data. If you know that you store visual memory with eye movements to the upper left or right, place your material approximately 12 inches in front of you on a desk and 6 inches toward the left or right (depending upon your preference) and you will create visual memory images as well as episodic memory. As a student or instructor, creating visual memory may even be as simple as moving your text from the left side of the desk to the right side of the desk.

References:

Barron, R. A. (2002). *Essentials of Psychology*. (3d Ed.) Allyn & Bacon

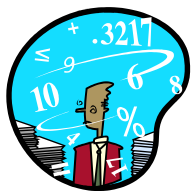
Grunebert, M. M. Morris, P., & Sykes, R. N. (1988). *Practical aspects of memory: Current research and issues* (Vols 1 & 2). Chichester, England: John Wiley & Sons.

Haberlandt, K. (1999). *Human memory: Exploration and application*. Boston: Allyn & Bacon.

Jenkins, J. G., & Dallenbach, K. M. (1924). Obliviscence during sleep and waking. *American Journal of Psychology*, 35, 605-612.

Tulving, E., & Psotka, L. (1971). Retroactive inhibition in free recall: Inaccessibility of information available in the memory store. *Journal of Experimental Psychology*, 87, 1-8.

Tulving, E., & Watkins, M. J. (1973). Continuity between recall and recognition. *American Journal of Psychology*, 86, 739-748.



Evaluation Tips

Instructional System Development – Evaluation Phase

by Donald Clark, copyright 1995, revised 1996 and 1997

Evaluating

This phase is ongoing throughout the entire process. That is, it is performed during the analysis, design, development, and implementation phases. It is also performed after the learners return to their jobs. Its purpose is to collect and document learner performance in a training course, as well as on the job. The goal is to fix problems and make the system better, not to lay blame.

Evaluation is the process of determining the value and effectiveness of a learning program. It uses assessment and validation tools to provide data for the evaluation. Assessment is the measurement of the practical results of the training in the work environment; while validation determines if the objectives of the training goal were met.

Bramley and Newby identify five main purposes of evaluation:

1. Feedback - Linking learning outcomes to objectives and providing a form of quality control.
2. Control - Making links from training to organizational activities and to consider cost effectiveness.
3. Research - Determining the relationships between learning, training, and the transfer of training to the job.
4. Intervention - The results of the evaluation influence the context in which it is occurring.
5. Power games - Manipulating evaluative data for organizational politics.

There are several factors which influence the outcome of the evaluation:

- Many trainers do not fully understand what constitutes the evaluation of training.
- The nature and type of organization influences the scope and methods of evaluation.
- The conduct of the valuation is dependent on whether internal or external evaluators are used.
- Lack of personnel trained in evaluation methodology (evaluator's expertise).

- The availability of resources. This includes time, money, and personnel.

Not every training course needs a evaluation or the same type of evaluation. Before beginning an evaluation, several questions need to be answered:

- Should an evaluation be done - is it worth the time and effort?
- What is the purpose of the evaluation?
- What will be measured?
- How comprehensive should the evaluation be?
- Who has authority and responsibility?
- What is the source of the data and how will the data be collected and compiled?
- How should the data be analyzed and presented?

The types of questions that can be answered by an evaluation include:

- What impact did the training have on the organization?
- Was a return on our investment realized?
- Are the learners using their new techniques and processes back in the work environment?
- Did the program change attitudes, behaviors, or skills in a way that positively impacts business results?

Internal Evaluation

Everyone in the training system is charged with this step. Their focus should be on the instructional processes and the measurement of learning that was gained from the training program. The primary purpose is to determine whether the instructional development effort has accomplished what was intended. Enough data must be collected so that through time, the instruction can be improved based upon learner performance. If a large proportion of learners have trouble with the same segment of instruction, it is reasonable to conclude there is something wrong with the instruction.

External Evaluation

After the internal evaluation has been completed, one major question about the entire training program remains unanswered: Can the learners do the job for which they were trained? The entire training process is designed toward this end. If the graduates do not need what they were taught, or need additional instruction, this information needs to be fed back to the instructional designers.

The various instruments used to collect the data are questionnaires, surveys, interviews, observations, and tests. The model or methodology used to gather the data should be a specified step-by-step procedure. It should be carefully

designed and executed to ensure the data is accurate and valid.

Questionnaires are the least expensive procedure for external evaluations and can be used to collect large samples of graduate information. They should be tried out before used to ensure the recipients of the questionnaire understand the operation in a way the designer intended it to be. When designing questionnaires, keep in mind the most important feature is the guidance given for its completion. All instructions should be clearly stated...let nothing be taken for granted.

The main portion of the questionnaire consists of a list of task statements. The supervisor of the graduates may be asked to rate the graduates on their ability to perform each task listed on the questionnaire.

Revise System

Once any training deficiencies have been noted, the SAT process is repeated to correct that deficiency. This does not mean that the entire training program is rebuilt, just the portions which are not training the learners to standards.

The Four Levels of Training Evaluation

Perhaps the best known training methodology is *Kirkpatrick's Four Level Evaluation Model*:

Level 1 - Reaction As the word implies, evaluation at this level measures how those who participate in the program react to it. This level is often measured with attitude questionnaires (smile sheets) that are passed out after most training classes. This level measures one thing: the learner's perception (reaction) of the course.

They might be asked how well they liked the instructor's presentation techniques, how completely the topics were covered, how valuable they perceived each module of the program, or the relevance of the program content to their specific job. They might also be asked how they plan to use their new skills back on the job.

Learners are keenly aware of what they need to know to accomplish a task. If the training program fails to satisfy their needs, a determination should be made as to whether it's the fault of the program design or delivery.

This level is not indicative of the training's return on investment as it does not measure what new skills the learners have acquired or what they have learned will transfer back to their working environments. This has caused some evaluators to downplay its value. However, the interest, attention and motivation of the participants are critical to the success of any training program. People learn better when they react positively to the learning environment.

Level 2 - Learning This can be defined as the extent to which participants change attitudes, improve knowledge, and increase skill as a result of attending the program. It addresses the question: Did the participants learn anything? The learning evaluation requires post-testing to ascertain what skills were learned during the training. The post-testing is only valid when combined with pre-testing, so that you can differentiate between what they already knew prior to training and what they actually learned during the training program.

Measuring the learning that takes place in a training program is important in order to validate the learning

objectives. Evaluating the learning that has taken place is typically focuses on such questions as: What knowledge was acquired? What skills were developed or enhanced? What attitudes were changed?

Learning measurements can be implemented throughout the training program, using a variety of evaluation techniques. Measurements at level 2 might indicate that a program's instructional methods are effective or ineffective, but it will not prove if the newly acquired skills will be used back in the working environment.

Level 3 - Behavior The level of behavior is defined as the extent to which a change in behavior has occurred because the participants attended the training program. This evaluation involves testing the students' capabilities to perform learned skills back on the job. Level 3 evaluations can be performed formally (testing) or informally (observation). It determines if a behavior change has occurred by answering the question, "Do people use their newly acquired skills, attitudes, or knowledge on the job?"

It is important to measure behavior because the primary purpose of training is to improve results by changing behavior. New learning is no good to an organization unless the participants actually use the new skills, attitudes or knowledge in their work activities. Since level 3 measurements must take place after the learners have returned to their jobs, the actual Level 3 measurements will typically involve someone closely involved with the learner, such as a supervisor.

Although it takes a greater effort to collect this data than it does to collect data during training, its value is important to the training department and organization. Behavior data provides insight into the transfer of learning from the classroom to the work environment and the barriers encountered when attempting to implement the new techniques learned in the program.

Level 4 - Results This is defined as the final results that occurred because the participants attended the program: the ability to apply learned skills to new and unfamiliar situations. It measures the training effectiveness, "What impact has the training achieved?" This broad category is concerned with the impact of the program on the wider community (results). It addresses the key question: Is it working and yielding value for the organization? These impacts can include such items as monetary, efficiency, moral, teams, etc. Here we expand our thinking beyond the impact on the learners who participated in the training program and begin to ask what happens to the organization as a result of the training efforts.

While it is often difficult to isolate the results of a training program, it is usually possible to link training contributions to organizational improvements. Collecting, organizing and analyzing level 4 information can be difficult, time-consuming and more costly than the other three levels, but the results are often worthwhile when viewed in the full context of its value to the organization.

As we move from level 1 to level 4, the evaluation process becomes more difficult and time-consuming, although it provides information that is of increasingly significant value. Perhaps the most frequently used measurement is Level 1 because it is the easiest to measure. However, it provides the least valuable data. Measuring results that affect the organization is more difficult and is conducted less frequently, yet yields the most valuable information...whether or not the organization is receiving a return on its training investment.

Each level should be used to provide a cross set of data for measuring training program.

Item Analysis

One of the tools used in the evaluation process is an item analysis. It is used to "Test the Test". It ensures the testing instruments are measuring the required behaviors needed by the learners to perform a task to standard. When evaluating tests we need to ask the question: Do the scores on the test provide information that is really useful and accurate in evaluating student performance? The item analysis provides information about the reliability and validity of test items and learner performance. Item Analysis has two purposes: First, to identify defective test items and secondly, to indicate which materials the learners have and have not mastered, particularly what skills they lack and what material still causes them difficulty.

Item Analysis is performed by comparing the proportion of learners who pass an item in contrasting criterion groups. That is, for each question on a test, how many learners with the highest test scores answered the question correctly or incorrectly compared with the learners who had the lowest test scores?

The upper (U) and lower (L) criterion groups are selected from the extremes of the distribution. The use of very extreme groups, say 10 percent, would result in sharper a differentiation, but it would reduce the reliability of the results because of the small number of cases utilized. In a normal distribution, the optimum point at which these two conditions balance out is 27 percent.

NOTE: With the large and normally distributed samples used in the development of standardized tests, it is customary to work with the upper and lower 27 percent of the criterion distribution. Many of the tables used for the computation of item validity indices are based on the assumption that the "27 percent rule" has been followed. Also, if the total sample contains 370 cases, the U and L groups will each include exactly 100 cases, thus preventing the necessity of computing percentages. For this reason it is desirable in large test item analysis to use a sample of 370 persons.

Because item analysis is often done with small classroom size groups, a simple procedure will be used here. This simple analysis uses a percentage of 33 percent to divide the class in three groups, Upper (U), Middle (M), and Lower (L). An example will be used for this discussion. In a class of 30 students we have chosen 10 students (33 percent) with the highest scores and 10 students (33 percent) with the lowest scores. We now have three groups: U, M, and L. The test has 10 items in it.

Next, we tally the correct responses to each item given by the students in the three groups. This can easily be done by listing the item numbers in one column and prepare three other columns named U, M, L. As we go through each student's paper, we place a tally mark next to each item that was answered correctly. This is done for each of the ten test papers in the U group, then each of the ten test papers in the M group, and finally for each of the ten papers in the L group. The tallies are then counted and recorded for each group as shown in the

table below.

Simple Item Analysis Procedure: Number of Learners Giving Correct Response in Each Criterion Group

Item	U (10)	M (10)	L (10)	Difficulty (U + M + L)	Discrimination (U - L)
1	7	4	3	14	4
2*	10	10	9	29	1
3	8	6	4	18	4
4*	4	4	6	14	-2
5†	6	7	6	19	0
6	8	7	4	19	4
7*	3	0	0	3	3
8	10	7	5	22	5
9*	1	2	8	11	-7
10	8	5	3	16	5

A measure of item difficulty is obtained by adding the number passing each item in all three criterion groups (U + M + L) as shown in the fifth column. A rough index of the validity or discriminative value of each item is found by subtracting the number of persons answering it correctly in the L group from the number answering it correctly in the U group (L - U) as shown in the sixth column.

Reviewing the table reveals five test items (marked with an *) that require closer examination.

Item 2 show a low difficulty level. It might be too easy, having been passed by 29 out of 30 learners. If the test item is measuring a valid performance standard, then it could still be an excellent test item.

Item 4 shows a negative value. Apparently, something about the question or one of the distracters confused the U group, since a greater number of them marked it wrong than the L group. Some of the elements to look for are: wording of the question, double negatives, incorrect terms, distracters that could be consider right, or text that differs from the instructional material.

Item 5 shows a zero discriminative value. A test item of this nature with a good difficulty rating might still be a valid test item, but other factors should be checked. i.e. Was a large number of the U group missing from training when this point was taught? Was the L group given additional training that could also benefit the U group?

Item 7 show a high difficulty level. The training program should be checked to see if this point was sufficiently covered by the trainers or if a different type of learning presentation should be developed.

Item 9 shows a negative value. The high value of the negative number probably indicates a test item that was incorrectly keyed.

As you can see, the item analysis identifies deficiencies either in the test or in the instruction. Discussing questionable items with the class is often sufficient to diagnose the problem. In narrowing down the source of difficulty, it is often helpful to carry out further analysis of each test item. The table below shows the number of learners in the three groups who choose each option in answering the particular items. For brevity, only the first three test items are shown. The correct

answers are marked with an *.

Response Analysis of Individual Items

Item		Response Options			
		A	B	C	D
1	Upper	7*	2	1	0
	Middle	4*	4	2	0
	Lower	3*	3	4	0
2	Upper	0	0	10*	0
	Middle	0	0	10*	0
	Lower	0	0	9*	1
3	Upper	1	8*	0	1
	Middle	2	6*	1	1
	Lower	2	4*	2	2

This analysis could be done with just the items that were chosen for further examination, or the complete test. You might wonder why perform another analysis for the complete test if most of the test items proved valid in the first one. The answer is to see how well the distracters performed their job. To illustrate this, look at the distracters chosen for item 1. Although the first analysis showed this to be a valid test item, of the distracters chosen by the learners, only A and B we used. Nine learners choose distracter B, seven learners choose distracter C, while none choose distracter D. This distracter needs to be made more realistic or eliminated from the test item. This type of analysis helps us to further refine the testing instrument.

References

Brown, Frederick G (1971). *Measurement and Evaluation*. Itasca, Illinois: F.E. Peacock.

Kelly, T. L. (1939). *The Selection of Upper and Lower Groups for the Validation of Test Items*. *Journal of Educational Psychology*. Vol. 30, p.p. 17-24

Kirkpatrick, Donald, (1994). *Evaluating Training Programs*. San Francisco, CA: Berrett-Koehler Publishers, Inc. (NOTE: Donald L. Kirkpatrick is a HRD Hall of Fame member.)

Student Evaluation of Instruction – A Case for Validity of Instructional Rating Forms

By Christina R. Silence

Should I provide an entry-level student an opportunity to rate the effectiveness of my instruction? After all, what does he/she know? If you find this to be your opinion or find that you are placed in a position to defend the validity of instructional rating forms, then I invite you to read on.

Based upon information gathered by the Center for Teaching and Learning at the University of North Carolina at Chapel Hill (UNC-CH), research has been done and you may find the results to be surprising. Although the research was performed in the college environment, we can see quite a parallel between the populations. Generally entry-level service members will be the same age as the traditional college student. Career-level service members will be the same approximate age as the non-traditional college student. One who possesses a master's degree or a doctorate in the college environment may provide the instruction; nonetheless, he/she is a subject matter expert who is being critiqued by a younger, less-educated or less-skilled group. This is comparable to

formal school instructors, the subject matter experts, who teach entry level.

According to the UNC-CH publication, studies have shown that "students and faculty offer very similar responses when asked to rank aspects of teaching in terms of their relative importance." In addition, there has been a "significant correlation between instructors' self-ratings of their effectiveness and student evaluations." So student evaluations are valid, but are they reliable? Studies also indicate that student evaluations tend to yield consistent results that are stable over time.

Many will argue that instructional rating forms are affected by the personality of the instructor, popularity, gender of instructor or student, or the time that the class is offered. However, there has not been evidence to support this. Studies did find that the instructor's rank did seem to have some effect on the ratings. Take for instance, college-teaching assistants did not tend to receive as high of ratings as faculty. First year faculty did not tend to receive as high of ratings as the more experienced faculty.

Interestingly enough, though not necessarily surprising, there was a trend that ratings by students tended to be harsher as they had been in college longer. Freshmen were more likely than sophomores to rate the same instructor higher. The same may be true when we look at ratings by entry-level students versus career-level students. The more experience an individual has in an instructional environment, the more critical he/she may be. His/her criticisms may also be more constructive; therefore more helpful to the instructional program. Ideas and recommendations are often indicated on instructional rating forms.

Instructional rating forms offer the student's perspective of instructional effectiveness. Additionally, they are a source of data for individual instructors who wish to improve their instruction. Of course instructional rating forms should not be the only means of measuring instructional effectiveness, but they do provide a perspective that should not be ignored. The overall ratings provide the most appropriate information for decision-making, while the questions targeting specific skills or attitudes are more diagnostic in nature. No school or instructor, for that matter, is so good that there cannot be improvement.

GENERAL INFORMATION

www.ott.navy.mil - This website is what I like to call "one stop shopping". It's full of information and links to other DOD training sites as well as public education sites. Many links can be found to the following five broad categories:

- ✓ Human Performance Measurement
- ✓ Needs Assessment
- ✓ Training Design and Development
- ✓ Education Technology
- ✓ Distance Learning

www.isd.uga.edu/teaching_resources/index.htm - The University of Georgia provides a variety of teaching resources on its Office of Instructional Support and Development website.

<http://jamaica.u.arizona.edu/ic/edtech/strategy.htm> | -The University of Arizona offers this website to provide information on different teaching methods and learning strategies.

www.eric.ed.gov - **ERIC** is a national information system funded by the [U.S. Department of Education's Institute of Education Sciences](http://www.ed.gov) to provide access to education literature and resources.

www.chaminade.org/inspire/learnstl.htm - Simple diagnostics chart that may help identify dominant learning modalities/styles. This is not an adequate substitute for taking a learning style inventory.

SPICE UP THE CLASSROOM

www.puzzlemaker.com - Instructors can make there own puzzles with their material on this website.

www.puzz.com - offers IQ tests, **puzzles**, riddles, humor, high IQ society information, games, and more. These can be used to reengage students after a break.

<http://www.thiagi.com/games.html> - 66 free training games.